|  |
| --- |
| 산학협력 프로젝트 수행계획서 |

|  |  |  |  |
| --- | --- | --- | --- |
| **과제명** | 대규모 인공지능 언어모델을 이용한 서비스 개발 | | |
| **협력기관명** | 구글 | **과제멘토** | 정명훈 이사 |
| **책임교수** | 정설영 | **소속** | 컴퓨터학부 |
| **참여인원** | (총 6명) 기업체 1명, 참여교수 1명, 학부과정 4명 | | |
| **수행기간** | 2023.09.01.∼12.31.(4개월) | **유형** | 중기 |
| **추진배경** |  | | |
| 최근 몇 년 동안, 인공지능 및 자연어 처리 분야(NLP)에서의 혁신적인 발전으로 대규모 언어모델, 특히 Google의 Bard나 Open AI의 Chat GPT 등과 같은 모델들이 주목을 받고 있습니다. 이러한 모델들은 자연어 이해와 생성 능력을 탁월하게 갖추고 있으며, 다양한 분야에서 혁신적인 서비스의 개발 가능성을 제시하고 있습니다. 따라서 우리는 이러한 인공지능 언어모델을 활용하여 학생들이 평소에 느끼는 불편과 문제점을 해결해주는 서비스를 개발하고자 합니다.  현대 대학생들은 다양한 강의와 교과목 중에서 최적의 선택을 하기 위해 많은 정보를 필요로 합니다. 그러나 강의평가 정보를 수집하고 정리하기는 번거롭고 시간이 소요되며, 종종 정보의 양과 질에 대한 의문이 있습니다. 이러한 문제를 해결하고 학생들이 최상의 교육 경험을 얻을 수 있도록, 대규모 인공지능 언어모델을 활용하여 강의 추천 서비스를 개발하고자 합니다. | | | |
| **목표 및 내용** |  | | |
| * Bard, Chat GPT 등의 자연어 처리 모델을 활용하여 비즈니스 아이디어를 서비스로 구현하는 실습 * 대학생들을 위한 강의 추천 서비스 개발 * 사용자가 원하는 조건(족보 유무, 시험 난이도, 과제량 등)을 입력하면 해당 조건에 맞는 강의를 추천 * Everytime 플랫폼의 강의평을 크롤링하여 모델 학습 * 언어 이해 능력을 향상시키기 위한 학습 데이터 수집 및 정제 * 자연스러운 대화 흐름과 사용자 의도에 민감한 응답 메커니즘 구현 * 구글 코리아 오피스 방문 및 견학(옵션) * 웹 서비스 개발   ※ 필요시 도식, 그림 등 삽입 | | | |
| **기대효과** |  | | |
| * 언어모델을 기반으로 한 서비스를 개발로 사용자와의 상호작용을 개선하고, 더 효율적인 편리한 경험을 제공할 수 있습니다. * 학생들의 강의 선택과 대학 생활의 질을 향상시킬 수 있습니다. * 대학 강의평과 정보의 투명성이 증가됩니다. * 대학 강의의 정보검색 편의성이 향상됩니다. * 언어모델과 관련한 연구 및 개발 경험을 통해 새로운 지식과 기술 개발에 기여하여 학문적인 발전을 이룰 수 있습니다. | | | |

**1. 과제 목적 및 필요성**

현대 대학교에서 학생들은 다양한 강의를 수강하고, 이를 통해 학문적이고 전문적인 지식 습득을 목표로 합니다. 그러나 강의 선택에 실패하거나 부적절한 강의를 수강하는 경우, 학습 경험은 저하되고 학업 성취도에 영향을 미칩니다. 이 프로젝트의 주요 목적은 학생들이 최적의 강의를 선택하고 학습 경험을 향상시키는 데 도움을 주는 강의 추천 서비스를 제공하여 학업 성취도를 향상시키는 것입니다.

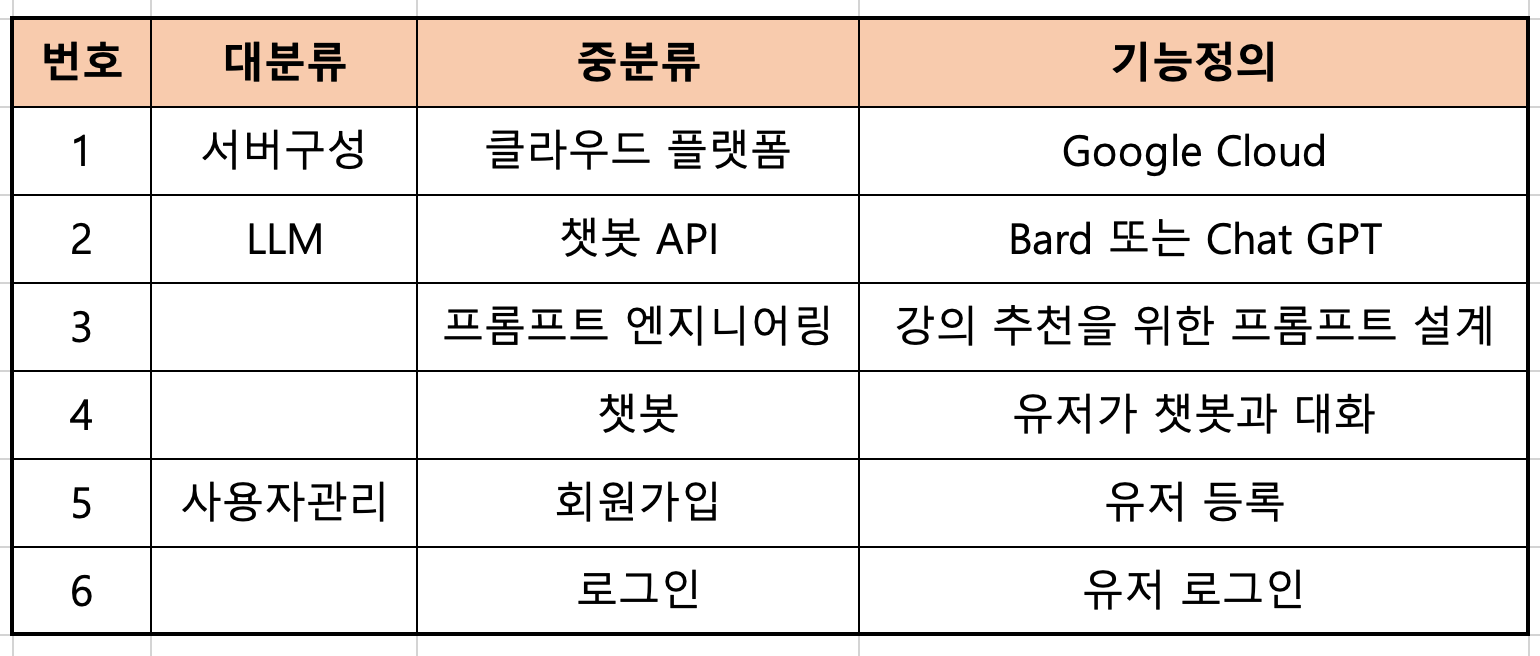
대부분의 학생들이 사용하는 플랫폼인 everytime의 강의평가 정보를 효과적으로 수집하고 활용함으로써, 학생들은 더 많은 정보를 획득하고 강의 및 교수진에 대한 투명성을 높일 수 있습니다. 또한, 학교와 학생들 간의 의사소통 강화를 통해 학교 운영의 효율성을 향상시키는 것도 이 프로젝트의 목적 중 하나입니다.

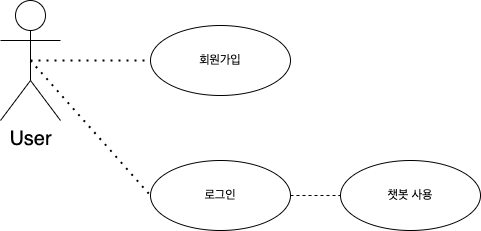
학생들은 주로 특정 강의를 선택할 때, 강의평가와 리뷰를 참고합니다. 그러나 이 정보를 수작업으로 수집하고 분석하기에는 시간과 노력이 많이 들며, 정보의 양과 질에 대한 의문이 있습니다. 이 프로젝트는 학생들이 최적의 강의를 선택하는 과정에 있어서 시간과 노력을 절약하고 최적의 선택을 할 수 있도록 도와줍니다.

※ 과제 내용을 충분히 전달할 수 있도록 목적 및 필요성 기술

(개략도, 모형도 등 첨부 가능)

**2. 과제 내용 및 추진 방법**

****

** **

* 데이터 수집 및 저장
* VSCode를 사용하여 대학 커뮤니티(예: Everytime)에서 강의평가 정보를 자동으로 크롤링 및 스크래핑하는 시스템을 개발합니다.
* 크롤링된 데이터를 구조화하고 MySQL 데이터베이스에 저장하여 향후 분석 및 활용을 위해 준비합니다.
* 대규모 인공지능 언어모델 학습
* 선택한 대규모 인공지능 언어모델(Large Language Model, LLM)을 활용하여 Colab을 사용하여 강의평가 데이터를 학습시킵니다.
* 데이터 전처리를 통해 모델이 데이터를 이해하고 관련 정보를 추출할 수 있도록 데이터를 가공합니다.
* 강의 추천 서비스 설계 및 구현
* 사용자 프롬프트에 따라 최적의 강의를 추천하는 서비스를 설계하고 구현합니다.
* ReactJS를 사용하여 사용자 인터페이스(UI)를 개발하고 사용자가 프롬프트를 입력하고 추천 결과를 확인할 수 있도록 합니다.
* 학습된 LLM을 활용하여 사용자의 요청에 따라 강의를 추천하고, 각각의 추천에 대한 설명 및 강의 정보를 제공합니다.
* 필요한 경우, 사용자가 추가적인 정보를 요청하거나 강의에 대한 리뷰를 열람할 수 있도록 서비스를 확장합니다.
* 테스트 및 피드백 수집
* 개발된 강의 추천 서비스를 학생들에게 테스트하고 피드백을 수집합니다.
* 사용자의 의견과 요구사항을 종합하여 서비스를 개선하고 사용자 경험을 향상시킵니다.
* 초기 사용자 테스트를 통해 시스템의 안정성을 확인하고 버그 및 문제점을 해결합니다.
* 서비스 배포 및 운영
* 개발된 강의 추천 서비스를 Google Cloud Platform을 사용하여 웹으로 배포합니다.
* 동계 계절학기 서비스를 학생들과 교수진에게 홍보하고 이용을 촉진합니다.
* 운영 단계에서 서버 호스팅 및 유지보수 작업을 수행하여 서비스의 안정성과 신뢰성을 유지합니다.

※ 과제 분석 내용 작성 (구현할 기술, 기능 등)

※ 추진을 위한 개발 환경, 툴, 장비, 재료 등의 활용 방안 및 절차 기술

**3. 과제 추진 일정 및 예산 활용 계획**

서버 운영비 : 0원

Google Cloud 비용 : 0원

Chat GPT API 비용 : 0원

논문 제출 비용 : 0원

오프라인 미팅 및 학회 이동교통비 : 0원



<https://www.notion.so/Gantt-chart-70a57f7d86684e49b31003e0bf39c468?pvs=4>

<과제 추진 일정 및 링크>

※ 과제 추진 일정은 Gantt 차트로 표현 (각 TASK 별 담당자 반드시 작성)

※ 주요 이벤트(논문발표, 대회참가 등)도 Gantt 차트 내 milestone으로 표현할 것

**4. 기대효과 및 활용방안**

이 과제의 성공적인 완수로 인해 학생들은 최적의 강의 선택에 도움을 받으며, 대학 커뮤니티는 강의와 교수진에 대한 투명성을 높이고 효율성을 향상시킬 수 있습니다.

※ 과제 수행 후 얻을 수 있는 효과 및 활용방안에 대하여 기술(교육적 효과, 상품화 방안 등)

**5. 예상되는 주요 과제성과**

프로젝트를 통해 개발된 강의 추천 서비스와 관련한 논문을 작성하여 학계와 업계에 결과를 발표할 예정입니다. 파라미터에 따른 프롬프트 모델의 성능을 주제로 논문을 작성할 예정입니다. 참석할 학회는 ICOT-2023과 한국정보과학회 총 두 개입니다. 이 논문은 학생들의 학습 경험과 대학 커뮤니티 혁신에 기여할 것으로 예상됩니다.

※ 논문발표(제출 학술대회, 일정 등), 특허출원, SW등록 등 (반드시 멘토/교수와 상의할 것)

**6. 참여인력(세부)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 지도교수 | 소속 | 컴퓨터학부 | | 성명 | 정설영 |
| 참여인력  (산업체) | 기업명 | 성명 | 직위 | 전화 | Email |
| 구글 | 정명훈 | 이사 | 010-7287-0851 | Javalove93@gmail.com |
| 과  제  참  여  학  생 | 소속(학과) | 학위과정  (성별) | 학번 | 성명 | 담당업무 |
| 컴퓨터학부 | 학사과정  (남) | 2021112751 | 김근찬 | 개발 |
| 컴퓨터학부 | 학사과정  (남) | 2017110157 | 신성한 | 개발 |
| 원예과학과 | 학사과정 (남) | 2018116035 | 신진철 | 조장, 개발 |
| 아동학부 | 학사과정  (남) | 2021115326 | 조정래 | 개발 |

※ 자유양식, 페이지 분량 제한 없음 (5~10페이지 권장)

**7. Github**

URL : <https://github.com/justgotothedesk/Comprehensive_Design_Project1>

김근찬 : mouse4432

신성한 : Hollyys

신진철 : justgotothedesk

조정래 : wjdfo